Strange uniform random variables

Douglas Rizzolo[*]

Department of Mathematics

University of Washington

January 31, 2013

Abstract

In probability theory, there is a tendency to treat one random variable with a given distribution as being just as good as any other. By and large this is fine because probability is (mostly) concerned with distributional properties of random variables. However, every now and again we are forced to deal with non-distributional properties. In this paper we investigate how different random variables with the same distribution can be. Specifically, we construct random variables that are all uniformly distributed on the unit interval, but that nonetheless have strikingly different properties.

In probability theory, there is a tendency to treat one random variable with a given distribution as being just as good as any other. By and large this is fine because probability is (mostly) concerned with distributional properties of random variables. However, every now and again we are forced to deal with non-distributional properties. For example, when working with Lévy processes, one often works with a cádlág version of the process even though the property of having cádlág paths is not distributional. Or, when working with Markov processes, one often works with the canonical version of the process on path space even though this is only one particular instance of the process. These examples highlight a natural question: given two random variables with the same distribution, how different can they really be?

In this note, we will mostly restrict ourselves to considering random variables that are uniformly distributed on $[0, 1]$. Throughout this note we let $\mathcal{B}[0, 1]$ be the Borel Ω-algebra on $[0, 1]$ and let λ be Lebesgue measure on $[0, 1]$ restricted to Borel sets. We will let $(\Omega, \mathbb{P}, \mathcal{F})$ denote an arbitrary, but not fixed, probability space. Recall that a measurable function $X : \Omega \to [0, 1]$ is called uniformly distributed if for all $A \in \mathcal{B}[0, 1]$ we have $\mathbb{P}(X^{-1}(A)) = \lambda(A)$. From this point on, we will adopt the usual probabilistic convention of using $\mathbb{P}(X \in A)$ instead of $\mathbb{P}(X^{-1}(A))$.

1

Let us begin our investigation by looking at the following scenario. Suppose that X and Y are uniformly distributed random variables (not necessarily defined on the same basic probability space) and that $h : [0, 1] \to \mathbb{R}$ is a function such that $h(X)$ and $h(Y)$ are Borel measurable. Does it follow that $h(X) =_d h(Y)$; that is, are $h(X)$ and $h(Y)$ necessarily equal in distribution? When h itself is measurable the answer is clearly yes, but the case when h is nonmeasurable is less clear. We will show that the answer is no, $h(X)$ and $h(Y)$ need not be equal in distribution. For the sake of full disclosure, we remark that this answer depends on using the Axiom of Choice. Later on we will comment on what can happen if the Axiom of Choice is not assumed.

Our first result is the following theorem.

Theorem 1. *Let μ and ν be two probability measures on $\mathbb{R}$. There exist uniformly distributed random variables $X : \Omega_1 \to [0, 1]$ and $Y : \Omega_2 \to [0, 1]$ and a function $h : [0, 1] \to \mathbb{R}$ such that $h(X)$ is measurable and has law μ and $h(Y)$ is measurable and has law ν.*

In fact, we can do even better (or worse, depending on your perspective).

Theorem 2. *Let I be an index set with cardinality at most that of the continuum and let $\{\mu_\alpha\}_{\alpha \in I}$ be a collection of probability measures on $\mathbb{R}$. There exists a family $\{X_\alpha\}_{\alpha \in I}$ of uniformly distributed random variables and a (single) function $h : [0, 1] \to \mathbb{R}$ such that, for all α, $h(X_\alpha)$ is measurable and has law μ_α.*

Another way to look at the difference between two random variables is to look at the set of values they take. More precisely, suppose that $X : \Omega_1 \to [0, 1]$ and $Y : \Omega_2 \to [0, 1]$ are uniformly distributed. What can we say about $X(\Omega_1) \cap Y(\Omega_2)$? Since $\lambda(\{x\}) = 0$ for all $x \in [0, 1]$ there is no particular $x \in [0, 1]$ that must be in this intersection, but how small can it really be? Could it be empty? This last question can be phrased more dramatically: is there a uniform random variable X such that we can remove all of the values it obtains from $[0, 1]$ and then build another uniform random variable that only takes values in what is left? Perhaps surprisingly, the answer is yes. We will prove the following strong version of the classical fact that a uniform random variable does not take any particular value.

Theorem 3. *There exist uniformly distributed random variables $X : \Omega_1 \to [0, 1]$ and $Y : \Omega_2 \to [0, 1]$ such that $X(\Omega_1) \cap Y(\Omega_2) = \emptyset$.*

Again, we can prove a stronger version.

Theorem 4. *Let I be an index set with cardinality at most that of the continuum. There exists a family $\{X_\alpha\}_{\alpha \in I}$ of uniformly distributed random variables defined on probability spaces $\{\Omega_\alpha\}_{\alpha \in I}$ such that $X_\alpha(\Omega_\alpha) \cap X_\beta(\Omega_\beta) = \emptyset$ whenever $\alpha \neq \beta$.*

We remark that Theorem 4 is quite similar to [8, Example 7.30], which proves essentially the same result using normally distributed random variables instead of uniformly distributed ones. Both results are essentially probabilistic reformulations of a classical theorem from measure theory that we will come to in due course (for the impatient reader, it is Theorem 6 below).

Let λ_* and λ^* be Lebesgue inner and outer measure respectively. We can also prove the following theorem.

Theorem 5. *There exists a family* $\{X_\alpha\}_{\alpha \in [0,1]}$ *of uniformly distributed random variables defined on probability spaces* $\{\Omega_\alpha\}_{\alpha \in [0,1]}$ *such that whenever* $\alpha \neq \beta$ *we have*

$$\lambda_*(X_\alpha(\Omega_\alpha) \cap X_\beta(\Omega_\beta)) = \min(\alpha, \beta) \quad and \quad \lambda^*(X_\alpha(\Omega_\alpha) \cap X_\beta(\Omega_\beta)) = \max(\alpha, \beta).$$

Theorems 1 and 2 can be obtained as consequences of Theorems 3 and 4 respectively.

Proof of Theorem 2. Let $\{X_\alpha\}_{\alpha \in I}$ of uniformly distributed random variables defined on probability spaces $\{\Omega_\alpha\}_{\alpha \in I}$ as in Theorem 4. For $\alpha \in I$, let F_α be the cumulative distribution function for μ_α and let F_α^{-1} be its right-continuous inverse. Define $h : [0,1] \to \mathbb{R}$ by

$$h(x) = \sum_{\alpha \in I} F_\alpha^{-1}(x) \mathbf{1}(x \in X_\alpha(\Omega_\alpha)).$$

Note that for each $x \in [0,1]$ at most one term in the sum is non-zero, so there is no need to worry about the fact that it is summing over a large index set. □

It remains to prove Theorem 4. The first step is to figure out exactly which subsets of $[0,1]$ can be the range of a uniform random variable. Suppose that $X : \Omega \to [0,1]$ is uniformly distributed and consider $X(\Omega)$. Observe that $X(\Omega)$ must intersect every set of positive measure since for all $A \in \mathcal{B}[0,1]$ we have $\mathbb{P}(X \in A) = \mathbb{P}[X \in (A \cap X(\Omega))]$. Thus $\lambda^*(X(\Omega)) = 1$. In fact, this condition is also sufficient for a subset of $[0,1]$ to be the range of a uniform random variable. Precisely,

Proposition 1. *A subset of* $[0,1]$ *is the range of some uniform random variable if and only if it has outer measure* 1.

Proof. From our discussion above, it remains to show that if $A \subseteq [0,1]$ has outer measure 1 then there is a uniform random variable with range A. We will construct such a random variable explicitly. Let $\mathcal{B}_A[0,1] = \{C \cap A : C \in \mathcal{B}[0,1]\}$. It is not hard to check that $\mathcal{B}_A$ is a σ-algebra on A.

Claim. *If* $B, C \in \mathcal{B}[0,1]$ *and* $B \cap A = C \cap A$ *then* $\lambda(B) = \lambda(C)$.

Proof of claim. In this case we have $A \cap B \cap C = A \cap B = A \cap C$, so that $B \setminus (B \cap C) \subseteq [0,1] \setminus A$. Since $\lambda^*(A) = 1$ this implies that $\lambda(B \setminus (B \cap C)) = 0$, so that $\lambda(B) = \lambda(B \cap C)$. Similarly, $\lambda(C) = \lambda(B \cap C)$, which proves the claim. □

We define $\lambda_A : \mathcal{B}_A[0,1] \to [0,1]$ by $\lambda_A(C \cap A) = \lambda(C)$. The claim shows that λ_A is well defined. It is straightforward to check that λ_A is a probability measure on $(A, \mathcal{B}_A[0,1])$. Define $X : (A, \lambda_A, \mathcal{B}_A[0,1]) \to ([0,1], \mathcal{B}[0,1])$ by $X(x) = x$. It follows immediately that X is uniformly distributed and has range equal to A. □

Proposition 1 reduces the proof of Theorem 4 to finding continuumly many pairwise disjoint subsets of $[0,1]$ each of which has outer measure 1. Fortunately for us, this is a classical result from measure theory due to Lusin and Sierpinksi,

Theorem 6 (Lusin and Sierpinski [5])**.** *There exists a partition of $[0, 1]$ into continuumly many pairwise disjoint sets, each of which has outer measure 1.*

The original reference for this Theorem is [5], and this paper has recently been reprinted in [6, p. 26]. See also [1] for a slightly more general version where $[0, 1]$ is replaced by an arbitrary closed set and the condition of having outer measure 1 is replaced by having full outer measure.

Something like the Axiom of Choice is needed to prove Theorem 6. To see this, let $(A_\alpha)_{\alpha \in I}$ be a partition of $[0, 1]$ into continuumly many sets such that $\lambda^*(A_\alpha) = 1$ for all $\alpha \in I$. Observe that for each α, A_α must be non-Lebesgue-measurable since we also have $\lambda^*([0, 1] \setminus A_\alpha) = 1$. The existence of non-measurable sets is connected, to some degree, with the Axiom of Choice. For example, there exists a model of the Zermelo-Frankel axioms of set theory (without Choice) in which the set of real numbers is a countable union of countable sets (see [3, Theorem 10.6]). In such a model, all sets are Borel, so nothing like what we have been discussing here can happen. Of course, it is also near impossible to do any analysis in such a model. Even if we allow enough Choice to do most analysis, things can go wrong. For instance (under some additional assumptions) there models of the Zermelo-Frankel axioms with the Axiom of Dependent Choice in which every set of reals is Lebesgue measurable (see [4, 7] for details).

Proofs of Theorems 4 and 5. For Theorem 4, let $(A_\alpha)_{\alpha \in I}$ be a partition of $[0, 1]$ as in Theorem 6 and for each $\alpha \in I$ let X_α be a uniform random variable with range A_α, as guaranteed to exist by Proposition 1. For Theorem 5, let $(A_\alpha)_{\alpha \in [0,1]}$ be a partition of $[0, 1]$ as in Theorem 6. Define $B_\alpha = A_\alpha \cup [0, \alpha]$ and each $\alpha \in [0, 1]$ let X_α be a uniform random variable with range B_α, again as guaranteed to exist by Proposition 1. $\square$

For the proof of Theorem 6 we refer the reader to the references mentioned above. We content ourselves to sketch the proof of the following weaker theorem,

Theorem 7. *There exists a partition of $[0, 1]$ into two sets, each of which has outer measure 1.*

For us, the advantage of proving Theorem 7 instead of Theorem 6 is that we can prove it using a minor variation of the standard construction of a non-measurable set. We follow the approach in [2, Section 16], to which we refer for a more detailed version of the proof.

Proof. Let ζ be an irrational number and let $B = \{n + m\zeta : n, m \in \mathbb{Z}\}$. Further define $B_e = \{n + m\zeta \in A : n \text{ is even}\}$ and $B_o = B \setminus B_e$. It is not difficult to see that B, B_e, and B_o are each dense subsets of $\mathbb{R}$. We define an equivalence relation on $\mathbb{R}$ by $x \sim y$ if and only if $x - y \in B$. Using the Axiom of Choice, we let $C \subseteq \mathbb{R}$ be a set containing exactly one representative of each equivalence class in $\mathbb{R}/\sim$. Let $M = C + B_e$. Note that the difference set $D(M) := \{m_1 - m_2 : m_1, m_2 \in M\}$ is disjoint from B_o. However, the difference set of any measurable set with positive Lebesgue measue contains an open interval about the origin. Since B_o is dense, it follows that M has no subsets of positive measure. Furthermore, we see that $\mathbb{R} \setminus M = M + 1$, so the complement of M also contains no subset of positive measure.

It follows that if we define $A_1 = M \cap [0, 1]$ and $A_2 = [0, 1] \setminus M$ then $\{A_1, A_2\}$ is a partition of $[0, 1]$ into two sets, each of which has outer measure 1. $\qquad\square$

References

[1] Alexander Abian. Partition of nondenumerable closed sets of reals. *Czechoslovak Math. J.*, 26(101)(2):207–210, 1976.

[2] Paul R. Halmos. *Measure theory*. Springer-Verlag, New York, 1974.

[3] Thomas J. Jech. *The axiom of choice*. North-Holland Publishing Co., Amsterdam, 1973. Studies in Logic and the Foundations of Mathematics, Vol. 75.

[4] Saharon Shelah. Can you take Solovay's inaccessible away? *Israel J. Math.*, 48(1):1–47, 1984.

[5] W. Sierpinski and N. Lusin. Sur une décompostion d'un intervalle en une infinité non dénumberable d'ensembles non measurables. *Comptes Rendus de l'Academie des Sciences*, pages 422–426, 1917.

[6] Yakov Sinai, editor. *Russian mathematicians in the 20th century*. World Scientific Publishing Co. Inc., River Edge, NJ, 2003.

[7] Robert M. Solovay. A model of set-theory in which every set of reals is Lebesgue measurable. *Ann. of Math. (2)*, 92:1–56, 1970.

[8] Gary L. Wise and Eric B. Hall. *Counterexamples in probability and real analysis*. The Clarendon Press Oxford University Press, New York, 1993.